MANGALORE UNIVERSITY

Final Synopsis

of the PhD Thesis to be submitted for the award of degree of Doctor of Philosophy in

Computer Science

TITLE OF THE WORK

Medical Images Analysis Using Deep Neural Networks

By

Mohammad Zaher Taljeh

Research Scholar

(Reg.No.MU/EXB/Ph.D/CR.111/Com.Sc./2022-23/E.25)

Under the Guidance of

Dr. B.H. Shekar,

Professor

Department of Post-Graduate Studies and Research in Computer Science,
Mangalore University,
Mangalagangotri-574 199

Sep 2025

Contents

0.1 Background . . . 1
0.2 Medical Image Preprocessing . . . 2
0.3 Medical Image Classification . . . 2
0.4 Medical Image Segmentation . . . 3
0.5 Various Image Acquisition Methods . . . 3
0.5.1 Ultrasound . . . 3
0.5.2 X-rays . . . 3
0.5.3 Magnetic Resonance Imaging(MRI) . . . 4
0.5.4 Mammograhy . . . 4
0.5.5 PET Scan . . . 4
0.5.6 Nuclear medicine imaging . . . 4
0.6 Literature Review . . . 4
0.6.1 Medical Image Segmentation . . . 4
0.6.2 Medical Image Classification . . . 6
0.6.3 Feature Extraction . . . 7
0.7 Limitations . . . 8
0.7.1 Data Limitations . . . 8
0.7.2 Technical Constraints . . . 8
0.7.3 Generalizability Issues . . . 8
0.7.4 Interpretation Challenges . . . 8
0.7.5 Practical Implementation Barriers . . . 9

1 Objectives **9**

2 Proposed Approaches **9**
2.1 Breast Cancer Segmentation Based on U-Net with Preprocessing Pipeline . 9
2.1.1 Abstract . . . 9
2.1.2 Proposed Methodology . . . 10
2.1.3 Dataset and Experimental Results . . . 10
2.2 SegNet Deep Architecture for Breast Tumor Segmentation . . . 13
2.2.1 Abstract . . . 13
2.2.2 Proposed Methodology . . . 13
2.2.3 Dataset and Experimental Results . . . 14
2.3 Transfer Learning Approach for Breast Ultrasound Image Segmentation . . 16
2.3.1 Abstract . . . 16
2.3.2 Proposed Methodology . . . 16
2.3.3 Dataset and Experimental Results . . . 18
2.4 U-Net + SegNet: An Ensemble Architecture for Breast Lesion Segmentation 19
2.4.1 Abstract . . . 19
2.4.2 Proposed Methodology . . . 20
2.4.3 Dataset and Experimental Results . . . 21
2.5 Attention U-Net with Transfer Learning for Breast and Skin Lesion Segmentation . . . 23
2.5.1 Abstract . . . 23
2.5.2 Methodology . . . 23
2.5.3 Dataset and Experimental Results . . . 24

2.6 Breast Cancer Classification Using AlexNet: A Dual-Input Approach Leveraging Ultrasound Images and Segmentation Masks 25
2.6.1 Abstract 25
2.6.2 Proposed Methodology 25
2.6.3 Dataset and Experimental Results 26

3 Conclusion & Future Works **27**

4 Publications **28**

Medical Images Analysis Using Deep Neural Networks

0.1 Background

Computer vision is a vital AI subfield focused on enabling machines to interpret and analyze visual data like images and videos. Using deep learning, it accomplishes tasks like object detection, image classification and segmentation. This technology underpins applications like autonomous driving, medical imaging and augmented reality. As a result, computer vision significantly advances AI's capacity to process and understand visual information Medical imaging is essential for modern healthcare, enabling accurate diagnosis, treatment planning and monitoring of various conditions. Techniques like ultrasound, MRI, CT, and X-ray provide detailed, non-invasive views of the body's internal structures, allowing clinicians to detect and assess abnormalities early. Advanced imaging analysis, often supported by AI, enhances diagnostic accuracy and helps personalize patient care, improving outcomes and reducing the need for invasive procedures. These modalities are crucial in areas like oncology, cardiology and neurology, making medical imaging a cornerstone of effective patient care. Interpreting medical images requires highly specialized expertise and radiologists and other imaging experts play a crucial role in diagnosing diseases through careful analysis of imaging data like X-rays, MRIs, and ultrasounds. However, there is a shortage of experienced radiologists worldwide, especially in rural or underserved areas, which can delay diagnoses and treatments.

AI in Computer-Aided Diagnosis (CAD) systems addresses this challenge by learning from vast amounts of imaging data labeled by expert radiologists. Through deep learning and pattern recognition, these systems can be trained to detect and classify abnormalities, mirroring the diagnostic insights of experienced doctors. By integrating AI, CAD systems not only support radiologists by flagging potential issues but also make high-quality diagnostic support accessible across healthcare settings. This approach amplifies the reach of expert knowledge, helping reduce diagnostic time and increasing accuracy even in regions where specialists are limited.

Medical image segmentation using deep learning is essential for accurate diagnosis and precise treatment planning, especially for conditions like breast cancer and skin lesions. In breast cancer detection, segmentation plays a vital role by identifying and isolating cancerous regions in mammograms, ultrasounds and MRIs. Precise segmentation helps radiologists determine the exact size, shape and spread of tumours, which is essential for early intervention and treatment planning. Similarly, in dermatology, accurate segmentation of skin lesions from dermoscopic images assists in distinguishing between benign and malignant lesions, enabling early detection of skin cancers like melanoma.

Deep learning-based segmentation models, such as U-Net and SegNet, excel at these tasks due to their ability to learn complex patterns directly from imaging data. By capturing both fine-grained details and broader contextual information, these models outperform

traditional methods, providing faster and more reliable results. This automation not only reduces the workload of medical professionals but also minimizes human error, enhancing diagnostic accuracy and potentially improving patient outcomes through early and precise detection of these life-threatening conditions. Medical imaging and analysis have become cornerstones of modern healthcare, fundamentally transforming diagnostic accuracy and treatment planning across medical specialties[25]. Advanced imaging modalities, including MRI, CT, and PET scans, provide clinicians with crucial insights into anatomical structures and physiological processes, enabling early disease detection and precise treatment monitoring[15]. The integration of artificial intelligence and deep learning algorithms has significantly enhanced image interpretation capabilities, reducing diagnostic time and improving accuracy in detecting abnormalities[31]. These technological advancements have particularly revolutionized oncology, where precise tumour visualization and volumetric measurements are essential for treatment planning and monitoring therapy response[17]. The emergence of 3D imaging reconstruction and real-time imaging guidance has dramatically improved surgical outcomes by providing surgeons with detailed anatomical information during procedures while also enabling minimally invasive approaches that reduce patient recovery time [18].

0.2 Medical Image Preprocessing

Preprocessing of raw medical image data is an essential step in medical image analysis, involving the enhancement and standardization of raw medical images to improve the reliability and accuracy of subsequent diagnostic or computational tasks. Techniques such as noise reduction, intensity normalization, contrast enhancement, artifact removal and spatial alignment address the inherent variability and complexity of medical imaging modalities like CT, MRI, and X-ray scans. Several investigations have highlighted the significant impact of integrating preprocessing techniques within deep learning pipelines, demonstrating substantial improvements in model performance compared to using the network independently. This integration not only enhances the quality of input data but also contributes to more robust and accurate medical image analysis [44]. Image preprocessing stands as a fundamental process in medical image analysis, critically enhancing the visual quality and diagnostic potential of medical images across various imaging modalities. Comprehensive preprocessing techniques, including contrast adjustment, noise reduction, intensity normalization and artifact removal, address intrinsic challenges inherent in medical imaging technologies like radiography, Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and ultrasound scans. Although these techniques do not inherently generate new diagnostic information, they strategically optimize image characteristics, unveiling subtle anatomical details and improving feature visibility for both human observers and advanced computational algorithms. By systematically mitigating technical variations, reducing statistical noise and standardizing imaging conditions, preprocessing plays a pivotal role in ensuring robust feature extraction, supporting advanced diagnostic interpretations, computer-aided detection systems and quantitative image analysis in clinical and research environments [52].

0.3 Medical Image Classification

Image classification involves categorizing and labeling groups of pixels within an image based on their features and predefined rules, making it a cornerstone in image processing. In medical imaging, classification plays a crucial role in identifying pathological patterns

or distinguishing between different anatomical structures, aiding significantly in diagnosis and treatment planning. The task of classification is broadly divided into supervised and unsupervised learning. Supervised classification relies on labeled datasets where ground truth is known, enabling models to learn the relationship between input features and expected outputs [34]. Conversely, unsupervised classification seeks to uncover hidden patterns or structures within unlabeled data [34]. Advanced techniques, including Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs), are widely employed in medical image classification. These methods leverage features such as texture, intensity and spatial patterns to classify images into categories, such as healthy tissue or various disease states. The emergence of deep learning, particularly CNNs, has revolutionized medical image classification by automating feature extraction and providing higher accuracy in tasks such as cancer detection, organ segmentation and disease diagnosis.

0.4 Medical Image Segmentation

Image segmentation is the process of dividing an image into distinct regions or objects by grouping neighboring pixels based on predefined similarity criteria. Many researchers have proposed various segmentation techniques; however, no universal method has been developed that can be applied across diverse applications. Segmentation divides an image into distinct regions based on specific properties, ensuring that each region remains homogeneous in terms of brightness, color, texture, and intensity. Generally, segmentation identifies areas of interest to facilitate the automation or enhancement of anatomical structure analysis. Image segmentation plays a crucial role in medical image analysis, as it is used to extract the Region of Interest (ROI) for further examination. Several feature extraction techniques exist for segmentation, including the Thresholding approach, Marker-Controlled Watershed segmentation, Region Growing segmentation, Random Walk method, Clustering-based segmentation and Deep Learning-based approaches[22].

0.5 Various Image Acquisition Methods

0.5.1 Ultrasound

A procedure that examines tissues and organs inside the body using high-energy sound waves is known as ultrasound imaging. A transducer generates and directs these sound waves toward the region of interest, with real-time results displayed on an integrated monitor, making it suitable for both dense and soft tissues [55]. This method is widely available, easily accessible, noninvasive, quick, and highly sensitive [55].

0.5.2 X-rays

X-rays are ionized forms of radiation that capture the image using rays. Images of the inside of your body are created using X-ray imaging. Different colours of black and white are used to depict the various sections of your body. This is because the amount of radiation absorbed by various tissues varies. Because calcium absorbs the most X-rays, bones appear white. Fat and other soft tissues are less absorbent and appear grey as a result. Because the lungs absorb the least amount of air, they appear black. The advantage of these applications lies in their ability to classify diseases based on information obtained from digital X-rays, producing images for immediate analysis. However, a drawback of

X-rays is that they may sometimes fail to accurately diagnose a disease due to image noise or blurring.[9].

0.5.3 Magnetic Resonance Imaging(MRI)

Magnetic Resonance Imaging (MRI) is a technique that uses radio waves and a powerful magnet connected to a computer to generate detailed images of internal body structures. These images help differentiate between healthy and diseased tissues. Compared to other imaging techniques such as CT scans and X-rays, MRI provides superior visualization of organs and soft tissues. It is recognized for its high accuracy in disease detection and its ability to produce detailed images of organs and other internal structures.[28].

0.5.4 Mammograhy

Mammography is an X-ray imaging technique used to detect early signs of breast cancer. It is considered the gold standard for breast cancer screening [41]. This method is highly effective in identifying cancer at an early stage, enabling timely treatment that may eliminate the need for chemotherapy. Additionally, it utilizes low doses of X-rays for imaging, ensuring minimal radiation exposure.

0.5.5 PET Scan

PET Scan is a process that involves injecting a small amount of radioactive glucose (sugar) into a vein and using a scanner to create detailed, digital images of locations inside the body where the glucose is taken up. The images can be used to locate cancer cells in the body because cancer cells often take up more glucose than normal cells also known as a positron emission tomography (PET) scan.[1] This method is distinguished by double the diagnostic clarity and easy and non disruptive. [2]

0.5.6 Nuclear medicine imaging

Nuclear medicine imaging is a technique that generates images by detecting radiation emitted from different parts of the body after a patient receives a radioactive tracer through injection, inhalation, or ingestion. This method is notable for its ability to produce clear images and detect diseases at an early stage.

0.6 Literature Review

This section provides an in-depth review of recent related studies, highlighting advancements in the field. We examine various methodologies, techniques, and findings from contemporary research to understand their contributions, limitations, and relevance to our work. By analyzing these studies, we aim to identify existing gaps and establish a foundation for our proposed approach.

0.6.1 Medical Image Segmentation

Cui et al. [10] introduced a segmentation method for breast ultrasound images, employing a combination of deep learning (SegNet) and an active contour model (ACM).

[1]https://siteman.wustl.edu/glossary/cdr000046140/

[2]https://www.independent imaging.com/the-pros-and-cons-of-pet-ct-scans/

Tong et al. [50] proposed an improved U-net based on mixed Attention Loss function (Improved U-net MALF) model to segment the breast tumors in ultrasound images.

Peiris et al. [38] introduced a semi-supervised image segmentation approach utilizing the principles of multi-view learning with the SegNet model. This method aims to tackle the challenge posed by the insufficient availability of pixel-level annotated data in medical images

Yap et al. [3] employed transfer learning to develop fully convolutional networks for breast mass segmentation, achieving superior results compared to traditional segmentation methods.

Hu et al. [16] investigated the applicability of FCN and U-Net in breast mass segmentation. They suggested that incorporating dilated convolutions in the deeper layers of FCNs could enhance segmentation performance. Dilated convolutions, compared to conventional ones, broaden the network's receptive field, facilitating the extraction of detailed spatial information.

Lyu et al. [30] introduced an enhanced method termed AMS-PAN (Improved Pyramid Attention Network) for breast ultrasound image segmentation. AMS-PAN is crafted to improve segmentation accuracy through the integration of Multi-Scale features and Attention mechanisms.

Meyer et al. [33] introduced an automated pipeline that utilizes deep convolutional neural networks to segment areas of COVID-19 infection in CT scans. They conducted an evaluation to assess how ensemble learning techniques such as bagging and augmenting impact the performance of the segmentation process.

Li et al. [24] proposed a method called Bagging Ensemble Deep Segmentation (BEDs) for training multiple U-Nets using fractional training data for segmenting dense nuclei in medical images.

Khaled et al. [54] proposed integrating three U-Net models into an ensemble technique, with each model using a distinct set of inputs for segmentation of the breast lesion from DCE-MRI.

El Adoui et al. [11] utilized two distinct and complete CNN encoder-decoder architectures SegNet and U-Net for segmenting breast lesion from DCE-MRI data and observed that U-Net outperformed SegNet.

Yan et al. [57] introduced an attention-improved U-Net model incorporating hybrid dilated convolution for the segmentation of breast ultrasound images.

Liu et al. [29] introduced a Multiscale Nonlocal Feature Extraction Network (MNFE-Net) comprising three essential components: a parallel encoder (PE), a multiscale feature module (MFM), and a global feature guidance module (GFGM). This network architecture was designed to attain accurate segmentation of breast abnormalities.

Corsi et al. [26] introduced a segmentation method that combines region- and edge-based information within a robust graph-based (RGB). This method efficiently segments the filtered image by employing a generative adversarial network (GAN) framework.

Xue et al. [56] proposed an innovative deep (CNN) architecture featuring specialized modules for detecting breast lesion boundaries along with a global guidance block. This network leverages integrated feature maps across multiple layers as informative guidance, allowing for the effective capture of long-range non local dependencies in both spatial and channel domains. The primary goal is to significantly improve the segmentation of breast ultrasound lesions.

Karimi et al. [21] employed transfer learning to train fully convolutional networks (FCNs) for medical image segmentation and random initialization to compare its effec-

tiveness. It also experimented with freezing the encoder and only training the decoder to assess the necessity of data-specific representations.

Sun et al. [47] introduced hierarchical weighting networks to evaluate the similarity between synthetic and real pixels, learning these weights at pixel, region and image levels to enhance model adaptation to the target domain.

0.6.2 Medical Image Classification

Veni and Manjula [53] propose the use of the Visual Geometry Group (VGG) network for brain image classification using transfer learning. They analyze four different VGG architectures and apply them to Magnetic Resonance Imaging (MRI) of the brain. The feature space is initially generated using multiple 3×3 convolutional filters, followed by pooling layers within each block to reduce dimensionality. These operations are repeated across subsequent blocks with varying numbers of convolutional filters. After a predetermined number of blocks, a fully connected layer with an activation function is used to classify the input MRI images. To evaluate performance, four VGG architectures with different depths—11, 13, 16, and 19 layers—were implemented. The results indicate that VGG architectures, when combined with transfer learning, exhibit strong potential for brain cancer classification.

S. Patil and Biradar [37] developed an optimized hybrid classifier for breast cancer detection. Their approach begins with noise removal from mammogram images using median filtering. The tumor is then segmented using an optimized region-growing algorithm called FC-CSO. For feature extraction, they utilize Gray-Level Co-occurrence Matrix (GLCM) and Gray-Level Run Length Matrix (GRLM) features. The extracted GLCM and GRLM outputs are fed into a Recurrent Neural Network (RNN) for classification, while the tumor-segmented binary image serves as input to a Convolutional Neural Network (CNN).

Muhammad et al. [45] have used optimized image processing and machine learning techniques to predict lung cancer. They examined the gathered images by using the multilevel brightness-preserving approach which actively examines each pixel, removes the noise and also increases the quality of the lung image. The effective features are then selected using hybrid spiral optimization intelligent-generalized rough set approach and those features are classified by using an ensemble classifier.

M. Salama and H. Aly [43] propose novel techniques for classifying Diffusion-Weighted Imaging (DWI) scans. Their approach utilizes two deep neural networks, ResNet50 and VGG16, with the final fully connected layer replaced by a Support Vector Machine (SVM) classifier to enhance classification performance. The dataset consists of 400 DWI images, with 211 patients diagnosed with prostate cancer and 189 without. To improve model generalization, data augmentation is applied by rotating each original image by 0°, 90°, 180°, and 270°.

Mustafa Koc et al. [23] propose a novel model that combines an ensemble feature generator using VGGNet (VGG16 and VGG19) with Neighborhood Component Analysis (NCA) for feature selection to enhance classification performance in prostate cancer detection. Their study is conducted on a dataset of 352 prostate MRI images.

Ruban Nersisson et al. [35] use CNN based on the YOLO architecture, named YOLO-CNN to extract features from the skin lesions. Traditional features like texture and colour extracted from the lesion region of the input images are concatenated with the features obtained from the CNN. After that, the concatenated features are supplied into a Fully

Connected Network for further processing.

Assiri et al. [4] proposed an ensemble classification approach based on a majority voting mechanism. Initially, they evaluated the performance of various state-of-the-art machine learning algorithms, including k-Nearest Neighbors (KNN), Simple Logistic Regression, Naïve Bayes, Multilayer Perceptron (MLP), Support Vector Machine (SVM) with Stochastic Gradient Descent (SGD), Random Decision Tree, Random Forest, and SVM with Sequential Minimal Optimization (SMO). These models were assessed using the Wisconsin Breast Cancer Dataset (WBCD), and the three best-performing classifiers were selected for the final ensemble classification.

Rabidas & Arif. [39] propose Local Photometric Attributes (LPA) for the characterization of mammographic masses as benign or malignant. LPA measures the local information over the optical density image which suppresses the background region and provides more details about the mass lesion. The introduced feature extraction module provides the Region of Interest (ROI).

KararAli et al. [2] developed a preprocessing image pipeline for this work. They removed hairs from the images, augmented the dataset and resized the images to meet the requirements of each model. By performing transfer learning on pre-trained ImageNet weights and fine-tuning the Convolutional Neural Networks, classification of images into target classes was carried out.

Domnguez and Thurnhofer-Hemsi [49] presents a deep learning system for skin cancer detection. Transfer learning was applied to five state-of-art convolutional neural networks to make both plain and hierarchical (with 2 levels) classifiers that are capable of distinguishing between seven types of moles and the work was carried out on the HAM10000 dataset.

0.6.3 Feature Extraction

Ojala et al. [36] presents a theoretically and computationally simple technique that is robust to grayscale variations and can effectively detect a wide range of rotated features. They also present a texture operator based on local binary patterns that are gray-scale and rotation invariant. They also present a texture operator based on local binary patterns that are gray-scale and rotation invariant. This gray-scale invariant operator incorporates a fixed set of rotation-invariant patterns to achieve rotation invariance. The work's key contribution is the recognition that certain local binary texture patterns known as uniform are essential features of local image texture, as well as the development of a generalized gray-scale and rotation invariant operator for detecting these 'uniform' patterns.

Lowe [27] proposed a new algorithm called Scale Invariant Feature Transform (SIFT) for feature extraction. At first it looks for critical spots in scale space that are the maxima or minima of a difference-of-Gaussian function. Each point produces a feature vector that describes the local image region sampled in the scale-space coordinate frame. The features are invariant to picture scaling, translation, and rotation, as well as somewhat invariant to changes in lighting and affine or three-dimensional projection. The SIFT keys derived from an image are used in a nearest-neighbor approach to indexing to distinguish prospective object models. sets of keys that agree on a possibility model pose are first identified through a Hough transform hash table, and then through a least squares fit to a final estimate of model parameters. When at least 3 keys agree on the model parameters with low residual, there is substantial evidence for the existence of the object.

HerbertBay. [6] present a new scale- and rotation-invariant interest point detector and

descriptor, coined SURF (Speeded Up Robust Features). It can be divided into three distinct steps. First, 'interest points' are chosen in the image at prominent locations such as corners, blobs, and T-junctions. After that, the neighbourhood of every interest point is represented by a feature vector. This descriptor has to be distinctive and at the same time, robust to noise, detection errors and geometric and photometric deformations. Finally, the descriptor vectors are matched between different images. SURF outperforms previously presented schemes with respect to repeatability, robustness and distinctiveness. This is done by relying on integral images for image convolutions; and by building on the strengths of the leading existing detectors and descriptors (specifically, using a Hessian matrix-based measure for the detector and a distribution-based descriptor).

0.7 Limitations

The application of deep learning in medical image analysis encounters several significant challenges:

0.7.1 Data Limitations

- Insufficient training data restricts the comprehensive learning and generalization of models.
- Limited diversity in medical image datasets potentially introduces bias and reduces model robustness.
- Small sample sizes significantly impact the performance and reliability of deep learning algorithms.
- Challenges in obtaining comprehensive and representative medical image collections.

0.7.2 Technical Constraints

- Substantial computational requirements for training and inference.
- Complex model architectures increasing computational complexity.
- Extended training times limiting rapid model development.
- Significant hardware resources necessary for effective implementation.

0.7.3 Generalizability Issues

- Difficulty in applying models consistently across different medical datasets.
- Variations in imaging equipment and protocols affecting model performance.
- Performance inconsistencies across diverse patient populations.
- Limited transferability of trained models between different clinical contexts.

0.7.4 Interpretation Challenges

- "Black box" nature of deep learning models obscuring decision-making processes.
- Difficulty in explaining model outcomes to medical professionals.

0.7.5 Practical Implementation Barriers

- Extensive clinical validation requirements.
- Significant challenges in integrating AI systems with existing medical infrastructure.
- Ethical concerns surrounding AI-assisted medical diagnosis and decision-making.

1 Objectives

Medical Image Processing is one of the exceptional achievements of computer science. It has significantly helped in discovering and early detection of many diseases and it had a significant role with artificial intelligence techniques in building cancer detection models that can assist in health centers and help medical experts affirm their medical verdicts without any obstruction.
Due to limited data size, high dimensionality and lack of efficient feature selection techniques existing image processing methods and artificial intelligence models need to be developed to increase their accuracy when detecting cancer because this will increase the chance of early treatment and cure. Cancer detection based on medical image processing and artificial intelligence is still an open area and needs a lot of work.

1. To design segmentation methods by exploring the intrinsic characteristics of medical images.
2. To develop preprocessing methods to improve deep learning model accuracy for cancer segmentation.
3. To design an ensemble model for segmentation and classification of skin and breast cancer.
4. To develop a method to classify medical image data for interpretation by domain experts.

2 Proposed Approaches

2.1 Breast Cancer Segmentation Based on U-Net with Preprocessing Pipeline

2.1.1 Abstract

One of the leading contributors of death for women around the world is breast cancer. Ultrasound is a common imaging method for detecting breast abnormalities. Because of the low contrast, high amount of shadows and blurred borders in breast ultrasound images, tumour segmentation is difficult. To overcome this, An approach based on the architecture of deep learning (U-net) with an image pre-processing pipeline has been proposed. We have proposed an improved methodology for the segmentation of Breast ultrasound images. To improve the image quality of a dataset of 647 breast ultrasound images(BUSI dataset), we have used a set of pre-processing techniques for speckle noise reduction and low contrast enhancement such as Contrast Limited Adaptive Histogram Equalization (CLAHE) and Gaussian Blur. The deep learning model (u-net) has then been trained and tested. To assess performance on the testing sets, Jaccard Index, Precision, Recall, and Dice Coefficient were calculated. The results were 81.63%, 89.87%, 89.75%

and 89.89%, respectively. In the same dataset, we compared the proposed method with three recent fully automatic segmentation methods and our method outperformed them in the evaluation metrics.

2.1.2 Proposed Methodology

1. **Gaussian Blur:** Noise is a type of disturbance that distorts the image. When there is noise in an image, the process of detecting boundaries becomes difficult because it disturbs all of the pixel values[46]. To decrease this noise we have used Gaussian Blur that is used to smooth an image and reduce the influence of noise [20]. It has the advantage of being extremely "smooth" and also circularly symmetric, which means that lines and edges in different directions are treated in the same way [12].

2. **Contrast Limited Adaptive Histogram Equalization (CLAHE):** CLAHE works on small portions of an image called tiles instead of the entire image. To eliminate artificial boundaries, the neighboring tiles are integrated by applying bilinear interpolation, which is one of the most widely used methods for improving image contrast [14]. This method remaps grayscale values from the input image to the output image, ensuring that the resulting histogram closely approximates a uniform distribution [42].

3. **Segmentation Architecture:**

   U-Net is a type of deep learning model used to segment images, where the goal is to quickly and accurately identify specific objects within an image. The U-Net model comprises the encoder and the decoder as its two major components.

   The encoder works by extracting high-level features from the inputted image. It works by applying a series of convolutional layers to the input image, which gradually decreases the feature maps' spatial dimensions while boosting several channels. The convolutional layers use learnable filters to scan over the inputted image and produce a set of output feature maps, where each map represents a different aspect of the inputted image. The output feature maps are then passed through a non-linear activation function, such as ReLU.

   At every layer of the encoder, the spatial resolution of the feature maps is decreased through pooling or stridden convolutions. This helps to gradually capture larger-scale features while also reducing The model's computational complexity. The output of the final layer in the encoder is a set of low-resolution, high-level feature maps that summarize the most important information in the input image. The decoder performs the opposite process of the encoder. It uses transposed convolutional layers to upsample the feature maps and gradually increase their spatial resolution. The decoder also includes skip connections that connect the corresponding encoder and decoder layers. These skip connections let the decoder use the low-level details and context from the encoder to refine the Map of segmentation. The final output of the decoder is a fully reconstructed image that has been generated from the encoded features and the skip connections[42].

2.1.3 Dataset and Experimental Results

This model was run on a machine with an Intel(R) Core(TM) i7-9700 CPU at 3.00GHz and 16 GB of Random-Access Memory, running Windows 11 64-bit Version 22H2. The

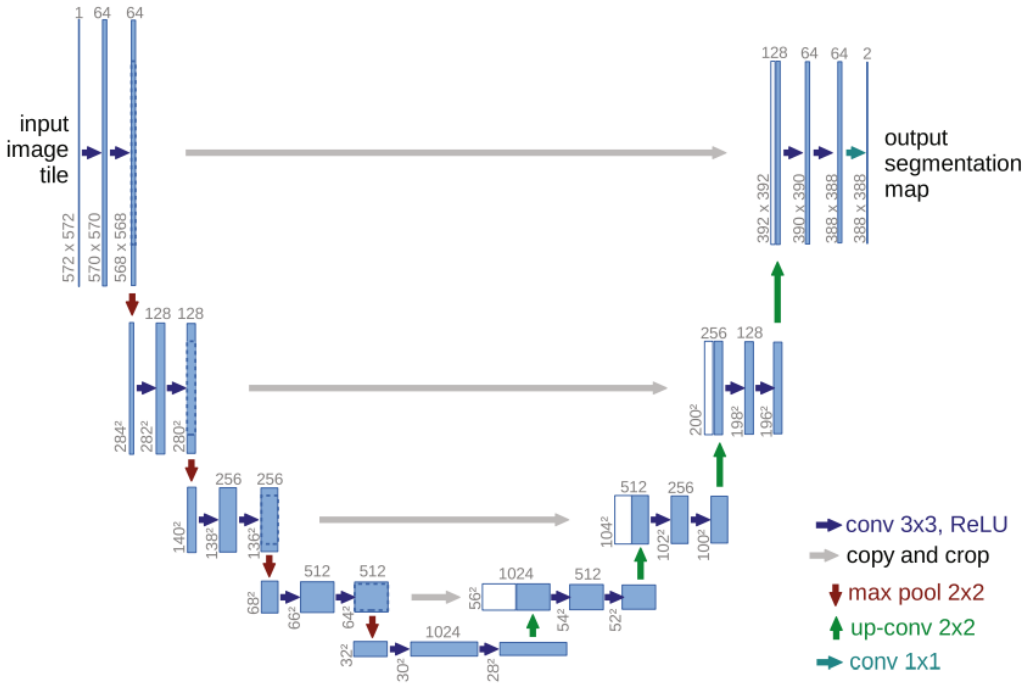

Figure 1: U-net architecture [42]

[42]

proposed model was trained and tested in Python through the Keras package with Tensorflow. The dataset used has been divided into three classes. 80% images from the dataset were taken for training, however, the remaining 20% were utilised for testing.

Dataset

Al-Dhabyani et al constructed this dataset (denotes as BUSI) [1]. The dataset contains 780 images obtained at the Baheya Hospital using two types of ultrasound equipment (LOGIQ E9 ultrasound and LOGIQ E9 Agile ultrasound system). The samples of this dataset are distributed as follows: 437 benign, 210 malignant, and 133 normal.

Evaluation Metrics

This study's evaluation metrics are the Jaccard Index, Precision, Recall, and Dice Coefficient.

- **True Positive (TP):**: Correct classification of an instance within the target category
- **False Positive (FP)**: Incorrect classification of an instance as belonging to the target category.
- **True Negative(TN)**: Correct classification of an instance as not belonging to the target category.
- **False Negative (FN)**: Incorrect classification of an instance as not belonging to the target category.

The performance metrics are discussed below:

- **Jaccard index(JI)**: Jaccard Index, also referred to as Intersection over Union (IoU) is utilized to determine the extent to which the ground(G) truth and prediction overlap (P). It is the most widely used and most effective in semantic segmentation.

$$Jaccardindex = \frac{G \cap P}{G \cup P} \tag{1}$$

where G and P symbolize the number of Ground truth and Prediction

- **Precision**: It is a metric that measures the quantity of accurate positive predictions.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

- **Recall**: It represents the ratio of positive samples that were correctly predicted to all positive samples.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

- **Dice coefficient**: It is used to gauge the similarity between the predicted and actual results.

$$DC = \frac{2 * TP}{2 * TP + FP + FN} \tag{4}$$

- **Accuracy**: The ratio of the number of correct predictions to the total number of predictions made.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

.

Results

In this work, we present a fully automated U-Net based method for segmenting breast tumours in which the model is trained after pre-processing through Gaussian Blur and CLAHE.

We conducted several experiments to assess our work with different settings. The proposed model was trained over 75 epochs as shown in Table 1 and the best performance results that the model scored were saved. Table 2 depicts the presented method's training and testing performance in terms of the Jaccard Index, Precision, Recall, and Dice Coefficient on the BUSI dataset. The proposed method demonstrates strong segmentation

Table 1: U-Net training parameters

Parameter	Learning-Rate (LR)	Epoch	Activation	Batch-size
Value	1e-2	75	Sigmoid	32

Table 2: The training and testing performance of the proposed method.

Method	Jaccard Index	Precision	Recall	Dice Coefficient
Training	84.90	91.06	90.96	91.08
Testing	81.63	89.87	89.75	89.89

performance, achieving a high Jaccard Index of 84.90 %, Precision of 91.06 %, Recall of 90.96 % and Dice Coefficient of 91.08 % during training. While testing metrics show a slight decline with Jaccard Index of 81.63 %, Precision of 89.87 %, Recall of 89.75 % and Dice Coefficient of 89.89 %, the model maintains reliable generalization. Overall, the method effectively balances precision and recall, ensuring accurate region identification with minimal false positives and negatives.

Through Table 3, we have conducted a comparison between the proposed method and the most advanced and up-to-date techniques.

Table 3: Comparative analysis of proposed method with state-of-the-art works.

Method	Jaccard Index	Precision	Recall	Dice Coefficient	Year
Hu et al[16].	-	-	-	82.69	2019
Byra et al[7].	-	-		70.9	2020
chen et al[8].	67.97	78.66	80.63	-	2022
Proposed Method	81.63	89.87	89.75	89.89	

2.2 SegNet Deep Architecture for Breast Tumor Segmentation

2.2.1 Abstract

Worldwide, breast cancer continues to be the primary cause of death for women, with ultrasound being a widely utilized imaging technique for detecting breast abnormalities. Tumour segmentation in ultrasound images for breast presents significant challenges due to issues such as low contrast, high levels of shadowing, and poorly defined borders. In this study, we propose an enhanced approach for breast ultrasound image segmentation by integrating an image pre-processing pipeline with a deep learning model based on SegNet. The preprocessing phase involved techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE), Gaussian blur and Gabor filtering to address issues of acoustic shadowing, speckle noise, and low contrast, thereby improving the quality of the dataset consisting of 647 breast ultrasound images (BUSI dataset). Following preprocessing, the SegNet model was trained and tested. Performance was evaluated using metrics such as Intersection over Union (IoU), Precision, Dice Coefficient and accuracy. The model achieved values of 94.65% for IoU, 79.65% for Precision, 74.66% for Dice Coefficient and 95.57% for Accuracy. Additionally, our proposed method was compared against four recent fully automatic segmentation techniques that work on the same dataset, demonstrating superior performance across all evaluation metrics.

2.2.2 Proposed Methodology

In this study, we employ a modified SegNet-based architecture [5] for the segmentation of breast ultrasound images from the BUSI dataset. The model is structured as an encoder-decoder network with enhancements tailored for improved segmentation performance. The

encoder begins with an input layer that accepts images of size 192×256×3. It comprises multiple convolutional layers with filters increasing in depth (64, 128, 256, 512), each followed by batch normalization and ReLU activation to stabilize training and improve feature learning. Each convolution block concludes with a max-pooling layer for downsampling, using indices that are stored for the decoder's upsampling operations. To capture high-level representations, two fully connected (dense) layers with 1024 units are added between the encoder and decoder, enhancing feature extraction before upsampling—a departure from the original SegNet architecture.

The decoder mirrors the encoder structure, using stored pooling indices to perform upsampling, which retains spatial information essential for precise segmentation. Each upsampling layer is followed by a series of Conv2DTranspose operations — used twice in each upsampling block to progressively refine feature representations and smooth the reconstruction—along with batch normalization and ReLU activation. The indices-based unspooling characteristic of SegNet is preserved, ensuring accurate reconstruction of key features from the encoder in the decoder. Additionally, our enhancements, including the use of Batch Normalization layers, increased model capacity and a modern optimization strategy, lead to improved segmentation accuracy and better performance metrics. The final layer employs a Conv2DTranspose operation followed by a sigmoid activation, generating a segmentation map that assigns pixel-wise probabilities. The output is reshaped to match the input image dimensions, facilitating a one-to-one mapping of segmentation results.

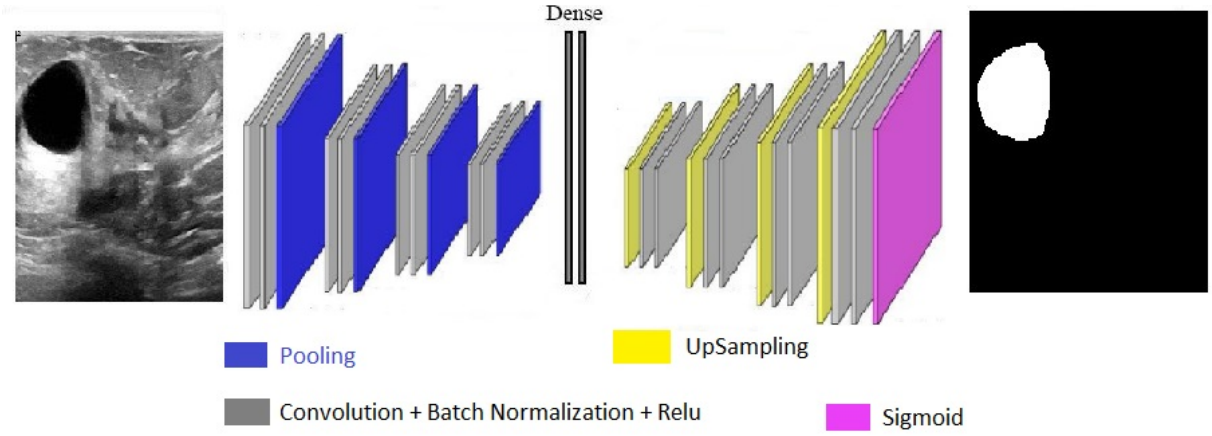

Figure 2: Architecture of the Proposed Model.

2.2.3 Dataset and Experimental Results

The model was executed on a PC operating Windows 11 64-bit Version 22H2, featuring an Intel® Core™ i7-9700 CPU clocked at 3.00 GHz and 16 GB of RAM. The implementation and evaluation of the proposed model were carried out using the Keras package with TensorFlow in Python. The dataset used in this study comprises three distinct classes, with 75% of the images allocated for training purposes.

Dataset:
The dataset utilized in this study, referred to as BUSI (Baheya Ultrasound for Breast

Cancer Imaging), was meticulously curated by Al-Dhabyani et al. [1]. The dataset encompasses a total of 780 images captured at Baheya Hospital, employing two distinct ultrasound equipment models, namely the LOGIQ E9 ultrasound and LOGIQ E9 Agile ultrasound system.

Results:
In this study, the SegNet model was utilized for the segmentation of breast ultrasonography images. The process involved several pre-processing steps, including image scaling, normalization, histogram equalization, and augmentation to enhance the dataset. To en-

Table 4: Performance metrics for the proposed model across training, validation, and testing phases.

Metric	**IOU**	**Precision**	**Dice Coefficient**	**Accuracy**
Training	99.06	97.34	95.76	99.33
Validation	95.67	88.38	80.52	96.92
Testing	94.65	79.65	74.66	95.57

Table 5: training parameters

Parameter	Learning-Rate (LR)	Epoch	Activation	Batch-size
Value	0.001	100	Sigmoid	18

sure robust model evaluation, a strict train-validation split was applied during the training phase, dedicating a specific portion of the enhanced dataset exclusively for model validation. This approach was critical in verifying the model's robustness and reliability. The model's effectiveness was assessed through a series of tests, varying the number of training iterations across 100 epochs systematically to identify optimal parameters. The outcomes are detailed in Table 4 & 5, which presents the training parameters and corresponding performance metrics. Table 6 provides a comparison of our method's performance against recent state-of-the-art techniques on the BUSI dataset, highlighting metrics such as Intersection over Union (IoU) Index, Precision, Recall and Accuracy.

Table 6: Comparison with Recent Studies of the Proposed Method

Method	**IOU**	**Precision**	**Dice Coefficient**	**Accuracy**
Chen et al.[8]	68.82	79.61	77.51	-
Tang et al.[48]	73.27	84.81	-	97.33
Lyu et al.[30]	68.53	83.50	80.71	97.13
Liu et al.[26]	77.72	-	85.49	98.18
Proposed Method	**94.65**	**79.65**	**74.66**	**95.57**

The results indicate that the model achieved high training accuracy with an IoU of 99.06%, Precision of 97.34% and Dice Coefficient of 95.76%. The validation metrics showed a slight decrease but remained robust with an IoU of 95.67%, Precision of 88.38%, and Dice Coefficient of 80.52%. The testing phase metrics were also strong, with an IoU of 94.65%, Precision of 79.65%, and Dice Coefficient of 74.66%, demonstrating consistent performance across different phases of evaluation.

The comparison shows that our method outperforms several recent techniques in terms of IoU and maintains competitive performance in Precision and Accuracy. Although our Dice Coefficient is slightly lower than some methods, our overall performance is strong, particularly in IoU and Accuracy.

2.3 Transfer Learning Approach for Breast Ultrasound Image Segmentation

2.3.1 Abstract

Breast cancer has been declared the leading cause of female mortality in the world. Ultrasound imaging continues to be the recommended method for detecting the presence of abnormalities in the breast tissue although tumour segmentation is difficult due to factors such as low contrast, high shadowing, and poorly defined borders. In this study, we propose an enhanced approach for breast ultrasound image segmentation by combining an extensive image pre-processing pipeline with a deep learning model based on SegNet, enhanced with transfer learning from MobileNet. Contrast Limited Adaptive Histogram Equalization (CLAHE), Gaussian blur, and Gabor filtering techniques have been put to use to reduce concerns of acoustic shadowing, speckle noise, and low contrast, improving the quality of the dataset, which consists of 647 breast ultrasound images (BUSI dataset). Additionally, data augmentation techniques were applied to further augment the dataset and improve model robustness. Using the pre-processed and augmented data, the SegNet model with MobileNet transfer learning was trained and evaluated. Model performance was assessed using metrics such as Intersection over Union (IoU), Precision, Dice Coefficient, and Accuracy. The proposed method was also compared against and found to be better than four recent fully automatic segmentation techniques on the same dataset.

2.3.2 Proposed Methodology

In this study, we developed an encoder-decoder model for segmenting breast ultrasound images, using MobileNet integrated into a SegNet-like backbone architecture. The proposed model combines the feature extraction capabilities of MobileNet with the segmentation structure of SegNet, focusing on effective feature extraction and rebuilding high-resolution segmentation maps.

The encoder-decoder model is designed to process and segment the input images effectively, and it includes the following parts:

Encoder: The encoder uses MobileNet, which is a lightweight and efficient model pre-trained on ImageNet. MobileNet is chosen because it can extract strong feature representations while keeping computational costs low. The required input shape for MobileNet is $256 \times 256 \times 3$ and it processes ultrasound images through several convolutional layers and depthwise separable convolutions.

The mathematical expression for feature extraction can be represented as:

$$E(I) = F(I) \quad \text{where } F \text{ is the feature extraction function of MobileNet}$$

MobileNet reduces the spatial dimensions while increasing the depth of feature maps, allowing it to capture important hierarchical features needed for segmentation. This architecture aligns with the principles of SegNet, which emphasizes both downsampling and upsampling pathways.

Dense Layer in the Bottleneck: Following MobileNet, the feature representation is passed through a dense (fully connected) layer, which further compresses the encoded

features into a compact vector in the latent space. This dense layer enables the model to learn abstract representations, capturing only the most relevant information for segmentation while reducing noise.

Decoder: The decoding process starts with upsampling the feature maps produced by MobileNet. UpSampling2D layers have been used to increase the spatial dimensions of these feature maps. After that, Conv2DTranspose layers are applied to further refine the segmentation results.

The decoding setup is structured as follows:

1. Initial Upsampling :

$$\text{Upsample}(X) \quad \text{increases the size of the feature map}$$

2. Convolutional Blocks : The model includes several Conv2DTranspose layers, each followed by Batch Normalization and ReLU activation to improve the feature representation. The configuration is:

- First, an UpSampling layer is applied, followed by three Conv2DTranspose layers, each with 512 filters. - Next, we perform more upsampling and convolution with 256 and 128 filters, respectively.

The mathematical expression for the decoding process is:

$$D(E(I)) = O \quad \text{where } D \text{ is the decoder function}$$

The final layers consist of a Conv2DTranspose with 64 filters, and then a single Conv2DTranspose layer with 1 filter to produce the final segmentation map. The output layer uses a sigmoid activation function to create binary segmentation masks, distinguishing between the foreground (lesion) and background.

To improve segmentation accuracy, the model uses indices pooling to retain spatial information from the encoder's max-pooling layers. These indices store the location of maximum values, which are used during upsampling in the decoder to reconstruct fine details in the segmentation mask. This approach allows the model to better capture boundaries and spatial relationships in the predicted output. The preserved indices guide accurate placement of features in the final segmented image.

Loss Function The model is trained using the Cross-Entropy Loss function, defined as:

$$L(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C} y_{i,c} \log(\hat{y}_{i,c})$$

where y represents the true labels, $\hat{y}$ denotes the predicted outputs, N is the total number of pixels, and C is the number of classes. This loss measures the difference between the predicted segmentation and the actual segmentation masks.

Optimization We optimize the model parameters using the Adam optimizer, updating the parameters according to:

$$\theta_t = \theta_{t-1} - \eta \frac{m_t}{\sqrt{v_t} + \epsilon}$$

where θ are the model parameters, η is the learning rate, m_t and v_t are the first and second moment estimates of the gradients, and ϵ is a small constant to avoid division by zero.

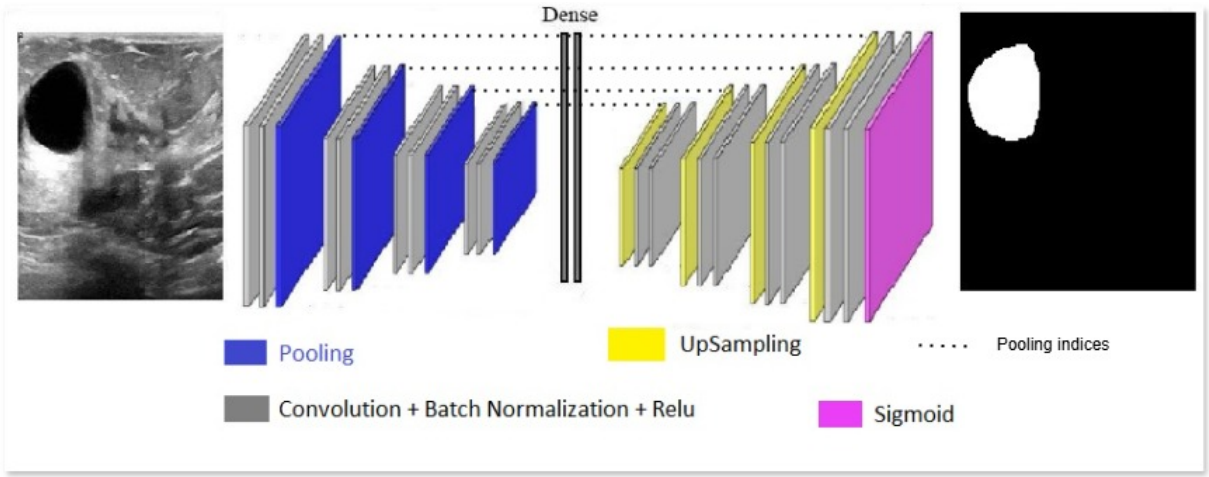

Figure 3: Architecture of the Proposed Model.

2.3.3 Dataset and Experimental Results

The model was executed on a PC operating Windows 11 64-bit Version 22H2, featuring an Intel® Core™ i7-9700 CPU clocked at 3.00 GHz and 16 GB of RAM. The implementation and evaluation of the proposed model were carried out using the Keras package with TensorFlow in Python. The dataset used in this study comprises three distinct classes, with 75% of the images allocated for training purposes and the remaining for testing.

Dataset:
The dataset utilized in this study, referred to as BUSI (Baheya Ultrasound for Breast Cancer Imaging), was compiled by Al-Dhabyani et al. [1]. It contains a total of 780 images captured at Baheya Hospital, employing two distinct ultrasound equipment models, namely the LOGIQ E9 ultrasound and LOGIQ E9 Agile ultrasound system.

Results:
In this work, we employed an encoder-decoder model using MobileNet as the encoder, enhanced with transfer learning, to segment breast ultrasound images. Our approach included several pre-processing steps, such as resizing the images to 256×256, normalization, histogram equalization, and data augmentation. These steps were carried out to enhance the dataset and improve the performance of the model.

Table 7: Performance metrics for the proposed model across training, validation, and testing phases.

Metric	IOU	Precision	Dice Coefficient	Accuracy
Training	96.95	95.90	95.72	99.40
Validation	80.22	89.89	83.45	97.51
Testing	72.17	79.78	73.92	95.68

To ensure robust model evaluation, a strict train-validation split was implemented during the training phase, reserving 25% of the enhanced dataset exclusively for validation purposes. This approach was essential for verifying the model's robustness and reliability. The model's effectiveness was systematically assessed by varying the number

Table 8: Training parameters

Parameter	Learning-Rate (LR)	Epoch	Activation	Batch-size
Value	0.001	100	Sigmoid	18

of training iterations over 100 epochs to determine the optimal parameters. The outcomes are summarized in Table 7 & 8, which details the training parameters and corresponding performance metrics. Table 9 provides a comparison of our method's performance against recent state-of-the-art techniques on the BUSI dataset, highlighting metrics such as Intersection over Union (IoU) Index, Precision, and Accuracy.

The results show that the model performed very well during training, achieving an IoU of 96.95%, a Dice Coefficient of 95.72%, precision of 95.90%, and accuracy of 99.40%. However, the performance metrics declined during testing, with an IoU of 72.17%, a Dice Coefficient of 73.92%, precision of 79.78%, and accuracy of 95.68%. This indicates that while the model was highly effective during training, its generalization to unseen data in the testing phase was somewhat reduced, though it still maintained strong performance overall.

The proposed method demonstrates superior performance compared to several recent approaches, particularly in terms of Intersection over Union (IoU), while also maintaining high precision and accuracy. Although the Dice Coefficient is marginally lower than some of the compared methods, the overall performance remains robust, with IoU and accuracy being especially strong indicators of the model's effectiveness.

Table 9: Comparison of the Proposed Method with Recent Studies.

Method Reference	**IOU**	**Precision**	**Dice Coefficient**	**Accuracy**
Chen et al.[8]	68.82	79.61	77.51	-
Lyu et al.[30]	68.53	83.50	80.71	97.13
Byra et al.[7]	-	-	0.646	94.40
Proposed Model	**72.17**	**79.78**	**73.92**	**95.68**

2.4 U-Net + SegNet: An Ensemble Architecture for Breast Lesion Segmentation

2.4.1 Abstract

Accurate segmentation of breast cancer lesions is vital for early detection, precise diagnosis, and effective treatment planning. In this research work, we propose an ensemble based segmentation approach that combines the power of the U-Net model and the SeqNet model for breast cancer segmentation. By leveraging the complementary strengths of these components, which capture both spatial and sequential information, we aim to improve the accuracy and reliability of breast cancer lesion segmentation, thereby facilitating better clinical decision-making and patient care. Experiments are conducted on BUSI dataset which is one of the standard datasets used by many researchers. A comparative analysis is also provided with the other approaches to justify the suitability of the proposed ensemble model for breast cancer segmentation task.

2.4.2 Proposed Methodology

The U-Net:
We build UNet model [40] with some customization options for the double convolution layer and input shape as follows:
Encoder:

It starts with an input layer (input) that takes the input images with shape `(192, 256, INPUT_CHANNELS)`. The encoder part consists of several conv Layers `(double_conv_layer)` followed by max pooling (MaxPooling2D) operations. The quantity of filters in the convolutional layers progressively increases, starting from filters and doubling at each stage (2x, 4x, 8x, 16x, 32x) compared to the previous layer.
Decoder:

The decoder part involves upsampling (UpSampling2D) layers. Subsequently, followed by the merging of these feature maps with their corresponding encoder counterparts. Each upsampling step is accompanied by a convolutional layer `(double_conv_layer)`, which reduces the number of filters (16x, 8x, 4x, 2x, 1x) compared to the concatenated input.
Final Layers:

After the decoder, the feature maps are further processed by a series of convolutional layers and activations. The last convolutional layer (Conv2D) with a 1x1 kernel produces the final output mask with the number of channels specified by `OUTPUT_MASK_CHANNELS`. To obtain the predicted masks, the output undergoes an activation function (sigmoid) application. The predicted masks are reshaped to the original image size using Reshape.
Model Compilation and Training:

The model is configured using the Adam optimizer, utilizing a learning rate of 0.003. The loss function employed is `jaccard_distance`. Multiple evaluation metrics are specified such as the Jaccard index, Dice coefficient, precision, recall, and accuracy. After training, the model is saved and returned. The U-Net architecture is frequently used for image segmentation tasks and has been customized for diverse applications, including medical image analysis.

The SegNet:
The Seqnet model, constructed as per[5], comprises 64 layers, excluding the final activation layer. The architecture and training configuration are detailed below.

The encoding layer begins with an input dimension of (192, 256, 3) and consists of consecutive convolutional layers, each followed by batch normalization and ReLU activation. These layers extract features at various abstraction levels, with max-pooling layers reducing spatial dimensions. The encoding layer further includes multiple sets of convolutional layers, progressively increasing filters from (64, 64) to (512, 512).

Following the encoding layers, two dense layers with 1024 units and ReLU activation capture global information. The decoding layer processes the output from dense layers, passing it through transpose convolutional layers, batch normalization, and ReLU activation. These transpose convolutional layers upsample feature maps to the original image size, employing multiple sets with decreasing filters from (512, 512) to (64, 1). The final layer uses a sigmoid activation function, yielding a segmentation mask reshaped to dimensions (192, 256).

For model compilation and training, stochastic gradient descent (SGD) is utilized with a learning rate of 0.001, momentum of 0.9, and a decay rate of 0.0005, excluding Nesterov

acceleration. The loss function is binary cross-entropy, and multiple evaluation metrics, including the Jaccard index, Dice coefficient, precision, recall, and accuracy, are specified. Training involves a set number of epochs and a batch size of 18, with validation data used to monitor performance. After training, the model is saved with the provided filename. This architecture and training setup is tailored for image segmentation tasks, to produce accurate segmentation masks for input images of dimensions (192, 256, 3).

Ensemble Bagging:

In this work, we leverage ensemble bagging as a core strategy to elevate the performance of our cancer segmentation models. Ensemble bagging is a technique that capitalizes on the synergy of multiple base models to enhance predictive accuracy and robustness. Our ensemble approach involves training independent instances of two renowned convolutional neural network architectures, the U-Net and SegNet models, with the overarching objective of refining the segmentation of cancerous regions within medical images. During the segmentation process, we employ a straightforward averaging strategy, whereby the predictions from these base models are element-wise averaged to generate a consolidated segmentation map. The advantages stemming from our ensemble bagging approach are highly pertinent to our specific cancer segmentation task. By embracing diversity among base models, our method significantly bolsters the robustness of cancer segmentation, rendering it less susceptible to overfitting and more dependable when analyzing medical images. .

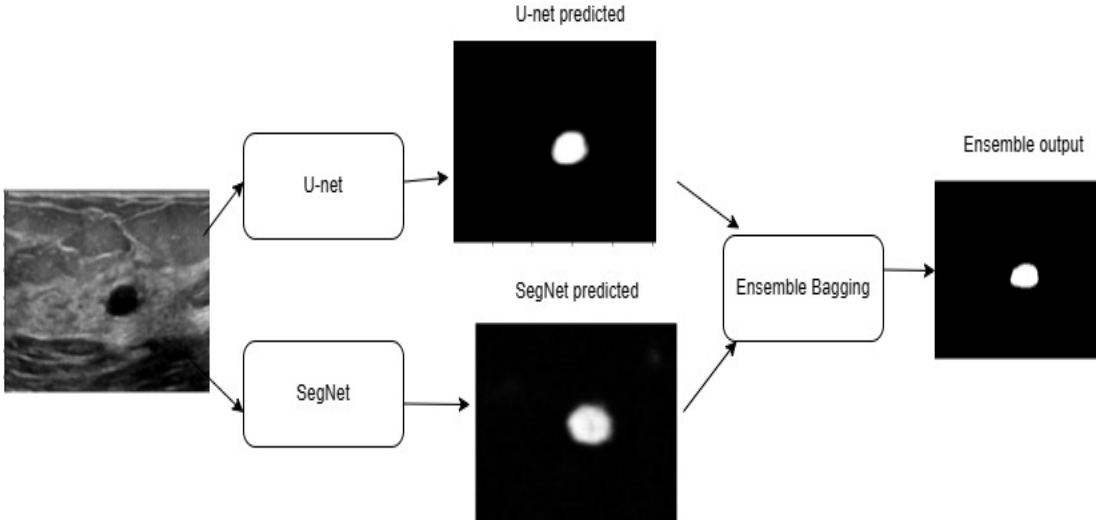

Figure 4

2.4.3 Dataset and Experimental Results

Dataset:

The dataset used for the research, referred to as BUSI (Breast Ultrasound Images), was

constructed by Al-Dhabyani et al. It consists of a total of 780 images that were obtained at the Baheya Hospital. The images were acquired using two types of ultrasound equipment, namely the LOGIQ E9 ultrasound and the LOGIQ E9 Agile ultrasound system[1].

The samples in the BUSI dataset are categorized into three classes based on their characteristics. The distribution of samples across these classes is as follows:

Benign: There are 437 images classified as benign. Benign breast lesions refer to non-cancerous or non-malignant abnormalities in breast tissue.

Malignant: There are 210 images classified as malignant. Malignant breast lesions represent cancerous or cancer-related abnormalities in breast tissue.

Normal: There are 133 images classified as normal. Normal breast images represent healthy breast tissue without any identified abnormalities.

Table 10: Modules training parameters

Parameter	**LR**	**Activation**	**Loss Function**	**Batch-size**	**Epoch**
U-net	0.003	Sigmoid	jaccard_distance	18	100
SegNet	0.001	Sigmoid	cross-entropy	18	100

Table 11: Comparison of the Proposed Method with Recent Studies

Method	**IOU**	**Dice**	**Precision**	**Recall**	**Accuracy**
Chen et al.[8]	68.82	77.51	79.61	81.10	-
Tang et al.[48]	73.27	-	84.81	84.26	97.33
Lyu et al.[30]	68.53	80.71	83.50	79.30	97.13
Liu et al.[26]	77.72	85.49	-	88.18	98.18
Proposed method	**95.87**	**72.57**	**84.96**	**70.46**	**96.05**

Results:

In this work, we present a ensemble bagging method that is based on the U-Net model and the SegNet model for segment breast ultrasound image, where the training was done after pre-processing the images (resize, normalize, and histogram equalization) and image augmentation. During the training phase, we adopted a rigorous train-validation split, allocating a portion of our augmented dataset exclusively for model validation. We conducted several experiments to evaluate our work, varying the number of training iterations for both the U-Net and SegNet models over 100 epochs to identify the optimal parameters, as presented in Table 10. Subsequently, we applied a simple post-processing technique using thresholding. The training time was 253,342.57 seconds and the testing time was 86.44 seconds.

We utilized Table 11 to compare the proposed approach with both the best and most recent methods available. This comparison enables us to assess the effectiveness and performance of our approach with existing techniques. We evaluated the following metrics for comparison: Intersection over Union, Dice Coefficient, Precision, Recall, and Accuracy.

2.5 Attention U-Net with Transfer Learning for Breast and Skin Lesion Segmentation

2.5.1 Abstract

Accurate and efficient segmentation of medical images is critical for the early detection and diagnosis of diseases such as breast cancer and skin lesions. In this study, we propose a deep learning-based segmentation approach utilizing Attention U-Net with transfer learning, leveraging VGG16 as the encoder to enhance feature extraction while maintaining computational efficiency. The model is trained and evaluated on two benchmark datasets: BUSI (Breast Ultrasound Images) for breast cancer segmentation and PH2 (Skin Lesion Images) for skin lesion segmentation. To further improve generalization and robustness, we apply data augmentation techniques and employ Dice loss combined with focal loss to address class imbalance. The experimental results as shown in Table 12 & 13, demonstrate that our method outperforms recent state-of-the-art approaches. The comparative analysis carried out has been listed in Table 14 & 15. On the BUSI dataset, the proposed model achieves an IoU of 89.41% and a Dice score of 72.11% on the test set, surpassing existing studies in terms of precision and recall. Similarly, on the PH2 dataset, the model attains an IoU of 94.26% and a Dice score of 92.15%, indicating superior segmentation accuracy. These results highlight the effectiveness of our approach in enhancing lesion boundary detection and segmentation performance, making it a promising tool for clinical applications.

2.5.2 Methodology

In this Work, we developed an advanced segmentation framework utilizing a modified attention-enhanced U-Net architecture with transfer learning capabilities for automated detection and segmentation of skin lesions and breast cancer imagery. The proposed architecture leverages a pre-trained VGG16 network (initialized with ImageNet weights) as the encoder backbone, enabling robust hierarchical feature extraction through its five convolutional blocks. Our decoder pathway implements a sophisticated attention mechanism through carefully designed attention gates, which compute spatial attention maps using 1×1 convolutions followed by softmax activation, allowing the model to dynamically focus on relevant anatomical regions during feature fusion. The network processes input images through multiple decoder blocks, each featuring transposed convolutions for upsampling, followed by batch normalization and ReLU activation functions. The skip connections from the encoder to decoder are enhanced with attention gates that help in selective feature refinement, particularly crucial for precise boundary delineation in medical imaging. To combat overfitting, we implemented spatial dropout with a rate of 0.5 and employed early stopping mechanisms monitoring validation loss. The model architecture includes multiple concatenation operations that merge attended feature maps from the encoder path with upsampled features from the decoder path, followed by convolutional blocks with batch normalization. The network was optimized using the Adam optimizer with a learning rate of 0.003, utilizing a comprehensive loss function combining Dice coefficient with custom components. Model performance was rigorously evaluated using multiple metrics including Jaccard coefficient, Dice coefficient, precision, recall, and accuracy, ensuring robust assessment of segmentation quality.

2.5.3 Dataset and Experimental Results

The experiments were conducted on a workstation with the following specifications: Processor: Intel(R) Core(TM) i7-9700 CPU @ 3.00GHz Memory: 16 GB RAM Operating System: Windows 11 64-bit, Version 22H2.

Datasets:
This study utilized two distinct datasets for image segmentation tasks:
BUSI: (Breast Ultrasound Images) Dataset: This dataset consists of breast ultrasound images, commonly used for breast cancer detection and segmentation tasks[1].
PH2 Dataset: The PH2 dataset is a dermoscopic image database used for skin lesion analysis and segmentation. It contains images of melanocytic lesions, including both benign nevi and malignant melanomas[32].

Result: We evaluated our model's performance using several standard metrics for image segmentation tasks. The results for both datasets are presented below:
PH2 Dataset Results:

Table 12: Model performance on PH2 dataset

Dataset	IOU	Dice Coefficient	Precision	Recall	Accuracy)	Loss
Train	95.64	93.86	96.32	91.62	96.16	4.36
Validation	95.40	93.44	94.41	92.56	95.85	4.60
Test	94.26	92.15	92.03	92.39	95.10	5.74

Table 13: Model performance on BUSI dataset

Dataset	IOU	Dice Coefficient	Precision	Recall	Accuracy	Loss
Train	92.54	94.52	98.62	96.79	99.05	1.38
Validation	91.99	73.39	94.46	88.69	95.99	5.45
Test	89.41	72.11	94.36	84.04	95.41	5.64

Table 14: Comparison of the Proposed Method on BUSI Dataset with Recent Studies

Method Reference	IOU	Dice	Precision	Recall	Accuracy
Chen et al.[8]	68.82	77.51	79.61	81.10	-
Tang et al.[48]	73.27	-	84.81	84.26	97.33
Lyu et al.[30]	68.53	80.71	83.50	79.30	97.13
Liu et al.[26]	77.72	85.49	-	88.18	98.18
Proposed Method for BUSI	**89.41**	**72.11**	**94.36**	**84.04**	**95.41**

The experimental results demonstrate the superior performance of our proposed method across multiple datasets and metrics. On the PH2 dataset, our method achieved remarkable results with an IOU of 94.26%, Dice score of 92.15%, Recall of 92.39%, and Accuracy of 95.10%, consistently outperforming recent state-of-the-art methods. Notably, our approach showed significant improvements over established methods like Valanarasu et al. (IOU: 82.77%) and Yin et al. (IOU: 87.84%). For the BUSI dataset, our model achieved

Table 15: Comparison of segmentation methods on PH2 dataset

Method	IOU	Dice Coefficient	Recall	Accuracy
Jianu et al.[19]	-	-	72.00	81.00
eVida.[13]	66.50	76.00	86.90	88.40
Yin et al.[58]	87.84	93.48		94.34
Valanarasu et al .[51]	82.77	90.51	90.82	91.58
Proposed Method for PH2	**94.26**	**92.15**	**92.39**	**95.10**

competitive results with an IOU of 89.41%, substantially surpassing previous approaches such as Chen et al. (IOU: 68.82%) and Liu et al. (IOU: 77.72%). The high precision score of 94.36% on the BUSI dataset indicates our model's strong ability to avoid false positives. While our Dice score on BUSI (72.11%) was lower than some benchmarks, the superior IOU and precision metrics suggest our model achieves better overall segmentation quality. The consistent performance across both datasets demonstrates the robustness and generalizability of our proposed method.

2.6 Breast Cancer Classification Using AlexNet: A Dual-Input Approach Leveraging Ultrasound Images and Segmentation Masks

2.6.1 Abstract

Breast cancer remains one of the leading causes of mortality among women, emphasizing the need for accurate and efficient diagnostic methods. This study presents a deep learning-based approach for breast cancer classification using ultrasound images and segmentation masks. The Breast Ultrasound Images (BUSI) dataset was utilized, incorporating both raw images and corresponding masks to enhance feature extraction. Due to the inherent class imbalance in the dataset, data augmentation techniques such as rotation, horizontal flipping, and color jitter were applied to generate synthetic samples and improve model generalization. Two deep learning architectures, AlexNet and DenseNet169, were trained and evaluated for the classification task. Experimental results demonstrate that AlexNet outperformed DenseNet169, achieving an accuracy of 99.14%, highlighting its effectiveness in breast cancer detection. The study underscores the potential of leveraging segmentation masks alongside ultrasound images to enhance classification performance, providing a reliable and automated tool for early diagnosis.

2.6.2 Proposed Methodology

This study employs a deep learning-based approach for breast cancer classification using ultrasound images and segmentation masks. The BUSI dataset was used, containing benign, malignant and normal ultrasound images. The dataset was split into training (315 benign, 152 malignant, 96 normal), validation (56 benign, 27 malignant, 17 normal) and test sets (66 benign, 31 malignant, 20 normal). To address class imbalance and enhance generalization, data augmentation techniques such as rotation, horizontal flipping, and color jitter were applied.

Two classification approaches were explored. The first used only ultrasound images with a pre-trained AlexNet model for transfer learning. Fine-tuned over 20 epochs with early stopping, it achieved 86.00% validation accuracy with a best validation loss of 0.4107.

Model	Train Loss	Train Acc.	Val Loss	Val Acc.	Train Time
AlexNet	0.0206	99.47	0.0359	99.00	17m 38s
DenseNet169	0.0299	99.47	0.0227	99.00	57m 16s

Table 16: Training and Validation Performance of AlexNet and DenseNet169

The second approach incorporated segmentation masks by overlaying them onto the original images and training a fine-tuned AlexNet model, significantly improving performance to 99.00% validation accuracy with a best validation loss of 0.0200. The final test evaluation achieved a perfect 100% accuracy across all classes.

The models were assessed using accuracy, loss function, precision, recall and F1-score. PyTorch was used for implementation, incorporating learning rate scheduling and weight decay. The results demonstrate that leveraging segmentation masks enhances classification performance, yielding superior diagnostic accuracy for breast cancer detection.

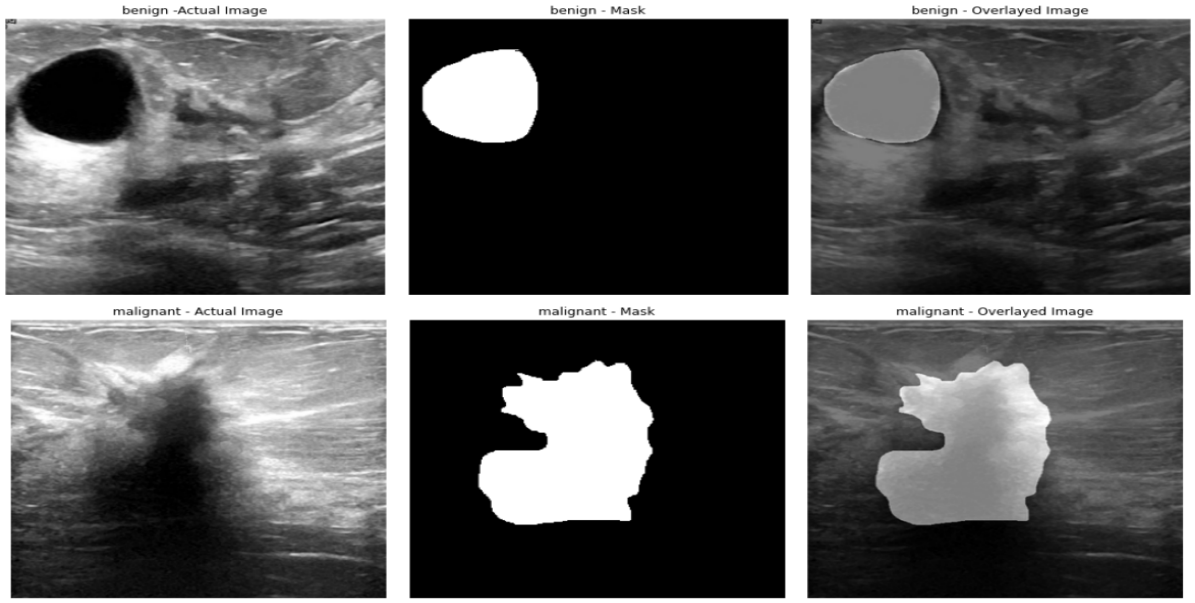

Figure 5: Sample images, corresponding masks and original images overlaid with corresponding masks.

2.6.3 Dataset and Experimental Results

The BUSI dataset was utilised for this research study. Figure 5 displays sample images from the dataset along with their corresponding masks. It also displays the images where the masks have been overlaid on the original ultrasound images. The results obtained for DenseNet169 model and the proposed AlexNet model have been listed in Table 16 and Table 17. Figure 6 lists the confusion matrix and heatmap obtained for the proposed AlexNet model. Some of the actual and predicted ultrasound images are shown in Figure 7.

Metric	AlexNet	DenseNet169
Accuracy	99.15	98.29
Precision	99.50	98.42
Recall	98.92	98.42
F1-score	99.20	98.42

Table 17: Test Performance Comparison of AlexNet and DenseNet169

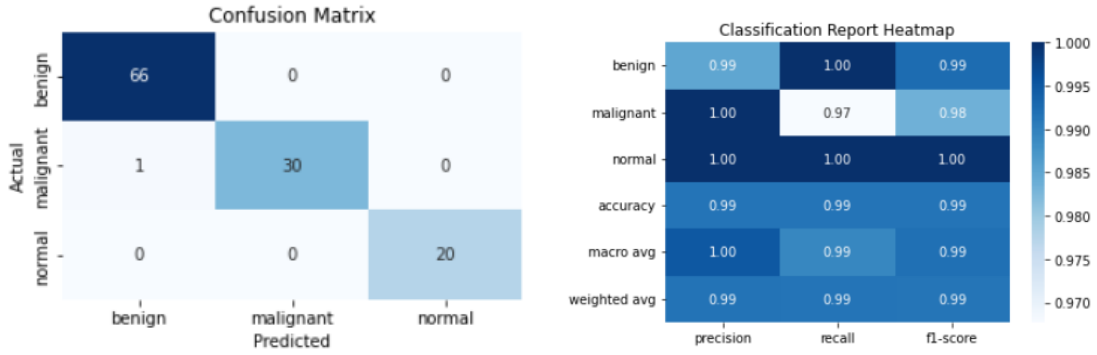

Figure 6: Confusion Matrix and Heatmap for the Proposed Model.

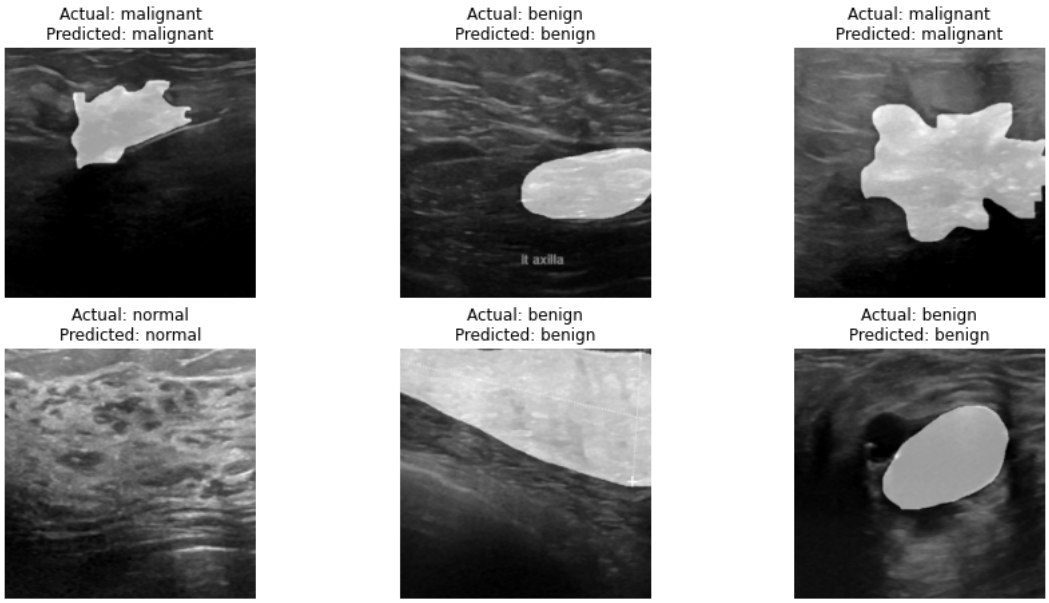

Figure 7: Actual and Predicted images obtained from the proposed model.

3 Conclusion & Future Works

This pre-thesis research investigates the application of deep neural networks for medical image analysis, with a focus on breast cancer and skin lesion segmentation and classification. A total of six deep learning-based approaches were explored to enhance segmentation

accuracy, classification performance, and overall diagnostic reliability. These methods address key challenges in medical imaging, such as noise reduction, tumor boundary preservation, and automated classification, to assist healthcare professionals in making accurate and efficient diagnoses.

For segmentation, U-Net with a preprocessing pipeline was implemented to improve image quality by enhancing contrast, reducing noise, and ensuring better feature extraction, leading to more precise segmentation of breast tumors. The SegNet deep architecture demonstrated its effectiveness in segmenting breast tumors by leveraging its encoder-decoder structure, ensuring better preservation of fine details and tumor boundaries. The transfer learning approach for breast ultrasound image segmentation utilized pre-trained models to extract high-level features from ultrasound images, improving generalization across different datasets. To further refine segmentation performance, an ensemble model combining U-Net and SegNet was introduced, leveraging the advantages of both architectures to achieve higher segmentation accuracy and robustness. Additionally, an Attention U-Net with transfer learning was developed for breast and skin lesion segmentation, incorporating attention mechanisms that enable the model to focus on critical lesion regions while filtering out background noise, ultimately enhancing segmentation precision and reducing false positives.

Beyond segmentation, a breast cancer classification approach using AlexNet with a dual-input strategy was proposed. This model integrates both ultrasound images and their corresponding segmentation masks, allowing the network to process raw image features alongside localized tumor information. This dual-input approach enhances classification accuracy by providing the model with both spatial and textural details of lesions, aiding in the differentiation between benign and malignant cases.

The findings from these studies indicate that deep learning-based segmentation and classification methods have significant potential in automating breast cancer diagnosis, reducing manual effort, and minimizing inter-observer variability in medical image interpretation. By integrating segmentation and classification, a more comprehensive and efficient workflow can be achieved, enabling early detection and improving patient outcomes. However, challenges remain, including computational efficiency, model interpretability, and the need for diverse and high-quality datasets to ensure robust generalization across various imaging conditions.

Future work will focus on further optimizing model architectures, improving segmentation mask quality, and incorporating multimodal imaging data, such as MRI and histopathology images, to enhance diagnostic precision. Additionally, exploring advanced deep learning techniques such as Vision Transformers (ViTs), self-supervised learning, and generative models can provide new insights into improving segmentation and classification accuracy. The integration of explainable AI (XAI) techniques will also be a key area of exploration, ensuring that deep learning models provide interpretable and trustworthy predictions for clinical decision-making. Through continuous advancements in deep learning, the proposed methodologies have the potential to significantly contribute to the field of medical image analysis, improving early diagnosis, treatment planning, and overall patient care.

4 Publications

- Shekar, B. H., & Mohammad Zaher Taljeh. **Enhancing Breast Cancer Lesion Segmentation with an Ensemble Of U-Net and SeqNet**. Presented at 9^th

International Symposium on Computer Vision and the Internet (VisionNet'23), September 20-December, 2023, Springer, organized by: PES University, Bangalore, India.

- Shekar, B. H., & Mohammad Zaher Taljeh. **Breast Cancer Segmentation Based on U-Net with Preprocessing Pipeline**. Presented at **International Conference on Sustainable Advanced Computing (ICSAC - 2024)**, March 22-23, Springer, organized by: CHRIST (Deemed to be University), Bangalore, India.

- Shekar, B. H., & Mohammad Zaher Taljeh. **SegNet Deep Architecture for Breast Tumor Segmentation**. Presented at **International Conference on Responsible Artificial Intelligence (ICRAI - 2024)**, December 16-17, 2024, Springer Nature, organized by: Mangalore University, Mangalore, India.

- Shekar, B. H., & Mohammad Zaher Taljeh. **Pre-Trained Encoder-Decoder Architecture for Breast Ultrasound Image Segmentation**. Accepted at 12^th **International Conference on Signal Processing and Integrated Networks (SPIN - 2025)**, February 20-21 2025, Springer, organized by: Amity University, Noida, India.

References

[1] Walid Al-Dhabyani, Mohammed Gomaa, Hussien Khaled, and Aly Fahmy. Dataset of breast ultrasound images. *Data in brief*, 28:104863, 2020.

[2] Karar Ali, Zaffar Ahmed Shaikh, Abdullah Ayub Khan, and Asif Ali Laghari. Multiclass skin cancer classification using efficientnets–a first step towards preventing skin cancer. *Neuroscience Informatics*, page 100034, 2021.

[3] Rania Almajalid, Juan Shan, Yaodong Du, and Ming Zhang. Development of a deep-learning-based method for breast ultrasound image segmentation. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 1103–1108. IEEE, 2018.

[4] Adel S Assiri, Saima Nazir, and Sergio A Velastin. Breast tumor classification using an ensemble machine learning method. *Journal of Imaging*, 6(6):39, 2020.

[5] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.

[6] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *European conference on computer vision*, pages 404–417. Springer, 2006.

[7] Michal Byra, Piotr Jarosik, Aleksandra Szubert, Michael Galperin, Haydee Ojeda-Fournier, Linda Olson, Mary O'Boyle, Christopher Comstock, and Michael Andre. Breast mass segmentation in ultrasound with selective kernel u-net convolutional neural network. *Biomedical Signal Processing and Control*, 61:102027, 2020.

[8] Gongping Chen, Lei Li, Yu Dai, Jianxun Zhang, and Moi Hoon Yap. Aau-net: an adaptive attention u-net for breast lesions segmentation in ultrasound images. *IEEE Transactions on Medical Imaging*, 2022.

[9] Akshay Trimbak Chikhalekar. Analysis of image processing for digital x-ray. *Int. Res. J. Eng. Technol.(IRJET) e-ISSN*, pages 2395–0056, 2016.

[10] Wenchao Cui, Dan Meng, Ke Lu, Yirong Wu, Zhihong Pan, Xiaolong Li, and Shuifa Sun. Automatic segmentation of ultrasound images using segnet and local nakagami distribution fitting model. *Biomedical Signal Processing and Control*, 81:104431, 2023.

[11] Mohammed El Adoui, Sidi Ahmed Mahmoudi, Mohamed Amine Larhmam, and Mohammed Benjelloun. Mri breast tumor segmentation using different encoder and decoder cnn architectures. *Computers*, 8(3):52, 2019.

[12] M Waltza Frederick and John WV Millerb. An efficient algorithm for gaussian blur using finite-state machines. In *E Conf. on Machine Vision Systems for Inspection and Metrology*, 1998.

[13] Jose Luis Garcia-Arroyo and Begonya Garcia-Zapirain. Segmentation of skin lesions in dermoscopy images using fuzzy classification of pixels and histogram thresholding. *Computer methods and programs in biomedicine*, 168:11–19, 2019.

[14] Rafael Gonzalez et al. R. woods digital image processing. *Addison–Wesely Publishing Company*, 1992.

[15] Ahmed Hosny, Chintan Parmar, John Quackenbush, Lawrence H Schwartz, and Hugo JWL Aerts. Artificial intelligence in radiology. *Nature Reviews Cancer*, 18(8):500–510, 2018.

[16] Yuzhou Hu, Yi Guo, Yuanyuan Wang, Jinhua Yu, Jiawei Li, Shichong Zhou, and Cai Chang. Automatic tumor segmentation in breast ultrasound images using a dilated fully convolutional network combined with an active contour model. *Medical physics*, 46(1):215–228, 2019.

[17] Saurabh Jha and Eric J Topol. Adapting to artificial intelligence: radiologists and pathologists as information specialists. *Jama*, 316(22):2353–2354, 2016.

[18] Jianmin Jiang, P Trundle, and Jinchang Ren. Medical image analysis with artificial neural networks. *Computerized Medical Imaging and Graphics*, 34(8):617–631, 2010.

[19] Serban Radu Stefan Jianu, Loretta Ichim, and Dan Popescu. Automatic diagnosis of skin cancer using neural networks. In *2019 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, pages 1–4. IEEE, 2019.

[20] Wen-Xiong Kang, Qing-Qiang Yang, and Run-Peng Liang. The comparative research on image segmentation algorithms. In *2009 First international workshop on education technology and computer science*, volume 2, pages 703–707. IEEE, 2009.

[21] Davood Karimi, Simon K Warfield, and Ali Gholipour. Transfer learning in medical image segmentation: New insights from analysis of the dynamics of model parameters and learned representations. *Artificial intelligence in medicine*, 116:102078, 2021.

[22] Charnpreet Kaur and Urvashi Garg. Artificial intelligence techniques for cancer detection in medical image processing: A review. *Materials Today: Proceedings*, 2021.

[23] Mustafa Koc, Suat Kamil Sut, Ihsan Serhatlioglu, Mehmet Baygin, and Turker Tuncer. Automatic prostate cancer detection model based on ensemble vggnet feature generation and nca feature selection using magnetic resonance images. *Multimedia Tools and Applications*, pages 1–20, 2022.

[24] Xing Li, Haichun Yang, Jiaxin He, Aadarsh Jha, Agnes B Fogo, Lee E Wheless, Shilin Zhao, and Yuankai Huo. Beds: bagging ensemble deep segmentation for nucleus segmentation with testing stage stain augmentation. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 659–662. IEEE, 2021.

[25] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *Medical image analysis*, 42:60–88, 2017.

[26] Guoqi Liu, Jiajia Wang, Dong Liu, and Baofang Chang. A multiscale nonlocal feature extraction network for breast lesion segmentation in ultrasound images. *IEEE Transactions on Instrumentation and Measurement*, 2023.

[27] David G Lowe. Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision*, volume 2, pages 1150–1157. Ieee, 1999.

[28] Alexander Selvikvåg Lundervold and Arvid Lundervold. An overview of deep learning in medical imaging focusing on mri. *Zeitschrift für Medizinische Physik*, 29(2):102–127, 2019.

[29] Yaozhong Luo, Longzhong Liu, Qinghua Huang, Xuelong Li, et al. A novel segmentation approach combining region-and edge-based information for ultrasound images. *BioMed research international*, 2017, 2017.

[30] Yuchao Lyu, Yinghao Xu, Xi Jiang, Jianing Liu, Xiaoyan Zhao, and Xijun Zhu. Amspan: Breast ultrasound image segmentation model combining attention mechanism and multi-scale features. *Biomedical Signal Processing and Control*, 81:104425, 2023.

[31] Morgan P McBee, Omer A Awan, Andrew T Colucci, Comeron W Ghobadi, Nadja Kadom, Akash P Kansagra, Srini Tridandapani, and William F Auffermann. Deep learning in radiology. *Academic radiology*, 25(11):1472–1480, 2018.

[32] Teresa Mendonça, M Celebi, T Mendonca, and J Marques. Ph2: A public database for the analysis of dermoscopic images. *Dermoscopy image analysis*, 2, 2015.

[33] Philip Meyer, Dominik Müller, Iñaki Soto-Rey, and Frank Kramer. Covid-19 image segmentation based on deep learning and ensemble learning. In *Public health and informatics*, pages 518–519. IOS Press, 2021.

[34] Eka Miranda, Mediana Aryuni, and E Irwansyah. A survey of medical image classification techniques. In *2016 International Conference on Information Management and Technology (ICIMTech)*, pages 56–61. IEEE, 2016.

[35] Ruban Nersisson, Tharun J Iyer, Alex Noel Joseph Raj, and Vijayarajan Rajangam. A dermoscopic skin lesion classification technique using yolo-cnn and traditional feature model. *Arabian Journal for Science and Engineering*, 46(10):9797–9808, 2021.

[36] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7):971–987, 2002.

[37] Rajeshwari S Patil and Nagashettappa Biradar. Automated mammogram breast cancer detection using the optimized combination of convolutional and recurrent neural network. *Evolutionary intelligence*, 14(4):1459–1474, 2021.

[38] Himashi Peiris, Zhaolin Chen, Gary Egan, and Mehrtash Harandi. Duo-segnet: adversarial dual-views for semi-supervised medical image segmentation. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part II 24*, pages 428–438. Springer, 2021.

[39] Rinku Rabidas and Wasim Arif. Characterization of mammographic masses based on local photometric attributes. *Multimedia Tools and Applications*, 79(29):21967–21985, 2020.

[40] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[41] Farahnaz Sadoughi, Zahra Kazemy, Farahnaz Hamedan, Leila Owji, Meysam Rahmanikatigari, and Tahere Talebi Azadboni. Artificial intelligence methods for the diagnosis of breast cancer by image processing: a review. *Breast Cancer: Targets and Therapy*, 10:219, 2018.

[42] Shelda Sajeev, Mariusz Bajger, and Gobert Lee. Segmentation of breast masses in local dense background using adaptive clip limit-clahe. In *2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pages 1–8. IEEE, 2015.

[43] Wessam M Salama and Moustafa H Aly. Prostate cancer detection based on deep convolutional neural networks and support vector machines: a novel concern level analysis. *Multimedia Tools and Applications*, 80(16):24995–25007, 2021.

[44] Massimo Salvi, U Rajendra Acharya, Filippo Molinari, and Kristen M Meiburger. The impact of pre-and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Computers in Biology and Medicine*, 128:104129, 2021.

[45] P Mohamed Shakeel, MA Burhanuddin, and Mohammad Ishak Desa. Automatic lung cancer detection from ct image using improved deep neural network and ensemble classifier. *Neural Computing and Applications*, pages 1–14, 2020.

[46] Pyushi Singhal, Akhilesh Verma, and Akash Garg. A study in finding effectiveness of gaussian blur filter over bilateral filter in natural scenes for graph based image segmentation. In *2017 4th international conference on advanced computing and communication systems (ICACCS)*, pages 1–6. IEEE, 2017.

[47] Ruoqi Sun, Xinge Zhu, Chongruo Wu, Chen Huang, Jianping Shi, and Lizhuang Ma. Not all areas are equal: Transfer learning for semantic segmentation via hierarchical region selection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4360–4369, 2019.

[48] Fenghe Tang, Lingtao Wang, Chunping Ning, Min Xian, and Jianrui Ding. Cmu-net: A strong convmixer-based medical ultrasound image segmentation network. *arXiv preprint arXiv:2210.13012*, 2022.

[49] Karl Thurnhofer-Hemsi and Enrique Domínguez. A convolutional neural network framework for accurate skin cancer detection. *Neural Processing Letters*, 53(5):3073–3093, 2021.

[50] Ying Tong, Yangyang Liu, Manxue Zhao, Lin Meng, and Jiachao Zhang. Improved u-net malf model for lesion segmentation in breast ultrasound images. *Biomedical Signal Processing and Control*, 68:102721, 2021.

[51] Jeya Maria Jose Valanarasu and Vishal M Patel. Unext: Mlp-based rapid medical image segmentation network. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference*, pages 23–33, Singapore, September 2022. Springer.

[52] P Vasuki, J Kanimozhi, and M Balkis Devi. A survey on image preprocessing techniques for diverse fields of medical imagery. In *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE)*, pages 1–6. IEEE, 2017.

[53] N Veni and J Manjula. High-performance visual geometric group deep learning architectures for mri brain tumor classification. *The Journal of Supercomputing*, pages 1–12, 2022.

[54] Joel Vidal, Joan C Vilanova, Robert Martí, et al. A u-net ensemble for breast lesion segmentation in dce mri. *Computers in Biology and Medicine*, 140:105093, 2022.

[55] Jessyca B Wagner. Artificial intelligence in medical imaging. *Radiologic technology*, 90(5):489–501, 2019.

[56] Cheng Xue, Lei Zhu, Huazhu Fu, Xiaowei Hu, Xiaomeng Li, Hai Zhang, and Pheng-Ann Heng. Global guidance network for breast lesion segmentation in ultrasound images. *Medical image analysis*, 70:101989, 2021.

[57] Yu Yan, Yangyang Liu, Yiyun Wu, Hong Zhang, Yameng Zhang, and Lin Meng. Accurate segmentation of breast tumors using ae u-net with hdc model in ultrasound images. *Biomedical Signal Processing and Control*, 72:103299, 2022.

[58] Wen Yin, Dongming Zhou, and Rencan Nie. Di-unet: dual-branch interactive u-net for skin cancer image segmentation. *Journal of Cancer Research and Clinical Oncology*, 149(17):15511–15524, 2023.

Research Scholar:	Research Supervisor:
Mohammad Zaher Taljeh, Research Scholar, Department of Computer Science, Mangalore University.	Dr. B.H. Shekar, Professor, Department of Computer Science, Mangalore University.